ANALYSIS OF NON-STATIONARY SIGNALS BY TIME FREQUENCY DISTRIBUTIONS

By

B HARSHAVARDHAN

DEPARTMENT OF ELECTRICAL ENGINEERING

Indian Institute of Technology Kanpur

MARCH, 2002

ANALYSIS OF NON-STATIONARY SIGNALS BY TIME FREQUENCY DISTRIBUTIONS

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of
Master of technology

by

B HARSHAVARDHAN

DEPARTMENT OF ELECTRICAL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY, KANPUR

MARCH, 2002

27-2-02

CERTIFICATE

It is certified that the work contained in the thesis entitled *Analysis of Non-stationary signals by Time-Frequency Distributions* by B Harshavardhan, has been carried out under my supervision, and this work has not been submitted elsewhere for the award of a degree

Pradip Sircar

March, 2002

Dr Pradip Sircar
Associate Professor
Department of Electrical Engineering
Indian Institute of Technology, Kanpur

ACKNOWLEDGEMENTS

I would like to express my deepest sense of gratitude and appreciation to *Dr Pradip Sircar* for his constant encouragement, valuable comments and constructive suggestions all through the completion of this work Special thanks are due to *Dr François Auger* of Boulevard de l'universite, France for his Time Frequency toolbox, which has been of ample use in the current work. I thank one and all who contributed to the success of this work directly or indirectly.

B Harshavardhan

ABSTRACT

A wide class of signals may be conveniently described in terms of time-dependent amplitude and frequency or sums of such amplitude and frequency-modulated components. However, the main problem is the numerical estimation of these time dependent characteristics Time–frequency representations offer a convenient setup to estimate these parameters which tend to concentrate the energy density in disjoint regions of time-frequency plane Scalogram, the modulus square of continuous wavelet transform, has been used as a time-frequency representation in the present work The local maxima of Scalogram, also known as ridges of the wavelet transform, contain crucial information on the characteristics of the signal Indeed, they mark the regions of the time–frequency plane where the signal concentrates most of its energy Complex AM and FM signal models are introduced and their suitability to model speech phonemes has been studied, the parameters of which are estimated using Scalogram

Contents

Chapter 1

1.1. Introduction 1
1.2. Motivation for taking up the problem 2
1.3 Organization of the Dissertation 4

Chapter 2

Time Frequency Distributions

2.1 Introduction 5
2.2. Scalogram 6
2.3. Instantaneous Frequency 8
2.4. Estimation of Instantaneous Frequency and Amplitudes 10
2.5. Finding the ridges

Chapter 3

Complex AM and FM Signal Models

3.1 Introduction 13
3.2. Complex AM signal model 14
3.2.1. Simulation Study 16
3.3. Complex FM signal model 23
3.3.1. Simulation Study 25

Chapter 4

Conclusions 32
Suggestions for further work 33

Bibliography 34

CHAPTER 1

1.1 Introduction

In many engineering applications such as speech analysis, speech synthesis, radar, sonar, telecommunications and engine diagnosis, the signals under consideration are known to be non-stationary, i e , their spectral contents vary with time Time–frequency analysis[1, 12], among other methods, was proposed to deal with such signals Time–frequency distributions (TFDs) are natural extensions of the Fourier transform They map a one-dimensional signal, to a two-dimensional quantity, a function of time and frequency A concept intimately related to TFDs is that of instantaneous frequency (IF) [2] In many situations, the IF characterizes important physical parameters of the signal Therefore, accurate and effective estimation of this quantity is of great importance

An important challenge of time-frequency analysis is the case of multi-component signals In general, real life signals are non-stationary and multi-component The time-frequency representation (also known as time-frequency distribution, TFD) characterizes such signals Several techniques have been proposed to reduce the effect of cross terms, which were inherent in any TFD The presence of cross terms obscures the time-frequency features of the signal TFDs concentrate the energy content of the signal in the time-frequency plane. Each such region in the time-frequency plane corresponds to a constituent component of the multi-component signal The idea is to get a distribution where each component can be separated from the rest which can be processed to extract the features of individual component Given a non-stationary signal which has it's energy localized in disjoint regions in the time-frequency plane, the main task is to develop procedures that manipulate components in the different regions independently In the case

of mono component signals, several ditributions[1] were proposed which characterize the signal But for the case of multi-component signals, several TFDs fail due to the presence of cross terms

For many years, the representation of a signal both in time and frequency domain has been of interest, especially when dealing with time-varying non-stationary signals like speech signals, seismic signals and EEG's The amplitude and frequency of such signals vary with time The most familiar representation is the spectrogram, which is, based on the assumption that for a short time basis, signals can be considered stationary Using this assumption spectrogram utilizes a short window whose length is chosen so that over the length of the window a signal can be considered stationary Then Fourier transform of the signal can be used to obtain the energy representation of the signal along the frequency direction at a given time, which corresponds, to the center of the window. The crucial drawback of this method is that the length of the window is directly related to the frequency resolution To increase the frequency resolution one must use a longer observation duration (i e , a longer window), which means that non-stationarities occurring during this interval are smeared in time and frequency The main disadvantage with the spectrogram is that it presents a constant time-frequency resolution and events separated by larger than the window length in time can not be effectively isolated.

1.2. Motivation for taking up the problem

A lot of work has been carried out in modelling non-stationary signals with varying degree of success The limitation of Fourier transform in analyzing non-stationary signals i e., those signals whose spectral description depends on time led to the tools such as time frequency distributions For example, Short Time Fourier Analysis

(STFT), Wigner Ville Distribution (WVD), Ambiguity Function (AF) etc , aim at describing time dependent spectral properties of signals, whereas time scale analysis, such as the wavelet transform aim at extracting localized contributions of signals which are labeled by scale parameter Wigner Ville Distribution (WVD) is very powerful for the analysis of mono component signals with linear frequency variation. The main drawback of WVD is that it is bilinear in nature, introducing cross terms in the time-frequency plane, which makes the transform difficult to interpret The wavelet transform is a time scale representation and is linear by definition The scale parameter, which is inversely proportional to frequency, can be adjusted to obtain a time-frequency representation comparable to STFT and WVD It has been shown that STFT and WVD are very much similar, in that both are outputs of a bank of filters whose impulse responses are $h(t)e^{j\omega t}$ *and* $\frac{1}{\sqrt{a}}g(t/a)$ respectively In the case of STFT a bank of filters $h(t)e^{j\omega t}$ can be considered as having constant bandwidth, since the window length is fixed in most of the cases, hence the spectral and temporal resolution is fixed in these cases. In the case of wavelet transform, however, a bank of filters $\frac{1}{\sqrt{a}}g(t/a)$ can be considered as a set of filters whose bandwidth is changing with frequency Since one can choose the scale parameter a which is inversely proportional to frequency, the wavelet transform provides high spectral resolution and low temporal resolution for low frequency Conversely for high frequencies, it provides high temporal resolution at the cost of spectral resolution Thus wavelet transform is a natural choice for the analysis of non-stationary signals

1.3. Organization of the Dissertation

This thesis is organized into a total of 4 chapters In chapter 2, time frequency distributions are reviewed, and Scalogram, a member of Time-frequency distributions of *affine class* is introduced It has been studied thoroughly and subsequently how it can be used to extract the instantaneous frequency and amplitude of an analytical signal is presented The method describes how ridges of the wavelet magnitude characterize completely the information present in an analytical signal The technique has been studied both for mono-component and multi-component cases In chapter 3, two novel models namely complex AM signal model and complex FM signal model are introduced. The Scalogram, which was introduced in the earlier chapter, has been used to estimate the parameters of the above models The suitability of the above models to fit voiced and unvoiced speech phonemes is carried out in the same chapter Chapter 4 concludes the whole work carried out and few suggestions are made for future work All the obtained results are attached right after their introduction in every chapter

CHAPTER 2

Time Frequency Distributions

2.1 Introduction

The power of standard Fourier analysis is that it allows the decomposition of a signal into individual frequency components and establishes the relative intensity of each component. The energy spectrum does not, however, tell us when those frequencies are occurred. It is in this context, the Short Time Fourier Transform (STFT), or spectrogram is introduced for the study of time-varying signals. However, there exists natural and man-made signals whose content is changing so rapidly that finding an appropriate short-time window is problematic since there may not be any time interval for which the signal is more or less stationary. Also, decreasing the time window so that one may locate events in time reduces the frequency resolution. Hence there is an inherent trade of between time and frequency resolution. Although the motivation behind time frequency distributions is to improve upon the spectrogram, the main idea is to find out means to analyze and classify time varying spectra. The basic idea that stemmed is to device a joint time-frequency function that will describe the energy density or intensity of a signal simultaneously in time and frequency.

2.2. Scalogram

The continuous wavelet transform (CWT) of a function x (t) ∈ $L^2(R)$ is defined as

$$T_x(a,b) = < x, \varphi_{a,b} > = \int_{-\infty}^{\infty} x(t)\varphi^*_{a,b}(t)\,dt \qquad 2\ 2\ 1$$

The function $\varphi_{a,b}$ represents a dilated and translated version of a basic function $\varphi(t)$ called *mother wavelet* *a* represents the scaling parameter of the wavelet, while *b* represents the dilation parameter The function $\varphi(t)$ can be called wavelet provided it satisfies the following conditions

- $\varphi(t)$ should integrate to zero $\int_{-\infty}^{\infty} \varphi(t)dt = 0$

this suggests that wavelet should have oscillatory or wavy appearance

- $\varphi(t)$ is square integrable or, equivalently, has finite energy $\int_{-\infty}^{\infty} |\varphi(t)|^2\, dt < \infty$

- $\varphi(t)$ should satisfy the admissibility condition

$$2\pi \int_{-\infty}^{\infty} \frac{|\hat{\varphi}(w)|}{|w|} dw < \infty \qquad 2.2\ 2$$

Conditions 1 and 2 suffice to define a wavelet, whereas admissibility condition is to be satisfied in formulating a simple inverse wavelet transform

The Scalogram of a signal x (t) is defined as the modulus square of the wavelet transform,

$$SC_x(a,b)=\left|T_x(a,b)\right|^2 \qquad 2\,2\,3$$

If $\varphi(t)$ satisfies the conditions mentioned in Eq (2 2 2), one can associate to it a time support interval T_φ, centered at a time t_o and a frequency support W_φ, centered at a frequency ω_φ.

$$\left.\begin{aligned} T_\varphi &= \sqrt{\int_{-\infty}^{\infty} (t-u)^2 \left|\varphi(t)\right|^2 dt} \\ W_\varphi &= \frac{1}{2\pi}\int_{-\infty}^{\hbar} (w-w_0)^2 \left|\hat{\varphi}(w)\right|^2 dw \end{aligned}\right\} \qquad 2.2\,4$$

Thus, $\varphi(t)$ has time-frequency support centered at $(t_\varphi,\omega_\varphi)$. It has been found that complex morlet wavelet out performs[10] among all other wavelets in localizing the energy of different components present in a multi-component signal Analytic wavelets are to be used to analyze the time evolution of frequency tones The analytic wavelet satisfies

$$\hat{\varphi}(w)=0 \quad , w\leq 0 \qquad 2\,2\,5$$

Complex morlet wavelet satisfies Eq (2 2 5) and it is defined as

$$\varphi(t)=e^{-\frac{t^2}{2\sigma^2}}e^{-jw_0 t} \qquad 2\,2\,6$$

which is a modulated gaussian function ω_0 is a frequency parameter controlling the number of oscillations of the wavelet within its gaussian envelope and σ controls the size of the envelope It has been found that the morlet wavelet has a compact support of [-4, 4] A Gaussian function is the most natural choice as it is the natural choice of a basic atom in time-frequency plane

2.3. Instantaneous Frequency

Consider the monochromatic wave, which can be unambiguously represented by

$$x(t) = a\cos(2\pi v_0 t) \qquad 2\ 3\ 1$$

Where the constants a and v_0 are the amplitude and frequency respectively It is quite tempting to extend this point of view to evolutionary situations, simply letting a vary in time, and by introducing an argument with a time varying derivative This would lead to

$$x(t) = a(t)\cos\phi(t) \qquad 2\ 3\ 2$$

Unfortunately the above expression is not unique There are infinitely many pairs (a (t), ϕ(t)) for the representation of x (t) A proper solution can be obtained to the above problem by taking the analytic signal of x (t), which would be given by

$$Z_x(t) = a(t)\exp(\phi(t)) \qquad 2.3\ 3$$

The analytic signal of x (t) is obtained by,

$$Z_x(t) = x(t) + jH\{x(t)\} = x(t) + \frac{j}{\pi} pv\int_{-\infty}^{\infty} \frac{x(s)}{t-s} ds \qquad 2\ 3\ 4$$

Where H denotes the Hilbert transform The complexified signal Z_x (t) is called the analytic signal The analytic signal admits a simple interpretation in frequency domain By definition, we get

$$Z_x(v) = X(v) + j(-j\,\mathrm{sgn}\,v)X(v) = 2U(v)X(v) \qquad 2\ 3\ 5$$

The real signal x (t) given by Eq (2 3 2) belonging to $L^2(R)$ with **a (t)** > 0 and **φ(t)** ∈ [0, 2π), t ∈ R, is said to be asymptotic if

$$\left|\frac{d\phi}{dt}\right| >> \left|\frac{1}{a}\frac{da}{dt}\right| \qquad 2.3.6$$

Which means that the oscillations coming from the phase term are much faster than the variation of amplitude If Eq (2 3 6) is satisfied, then we can approximate the analytic signal of Eq (2 3 2) to the form,

$$Z_x(t) \approx a(t)\,\mathrm{e}^{j\phi(t)} \qquad 2.3.7$$

The above analysis applies to each component of the analyzed multi-component signal

2.4 Estimation of Instantaneous Frequency and Amplitudes

Consider a mono component signal of the form

$$x(t) = A(t)\cos\phi(t) \qquad 2\,4\,1$$

Let $\varphi(t)$ represent an analytic wavelet, i e of the form,

$$\varphi(t) = g(t)e^{jw_0 t} \qquad 2\,4\,2$$

Where g (t) is the gaussian function The analytic signal of the mono component signal given by Eq (2 4 1) is given by

$$Z_x(t) = A(t)e^{j\phi(t)} \qquad 2\,4\,3$$

The CWT of the signal given by Eq (2 4 1) can be obtained as the inner product of the analytical signal and mother wavelet.

$$T_x(a,b) = < x, \varphi_{a,b} > = \frac{1}{2} < Z_x, e^{jw_0 t}\, g_{b,a} > \qquad 2\,4\,4$$

Under asymptotic condition[6] of Eq (2 3 6), it can be shown that

$$T_x(a,b) = \sqrt{\frac{a}{2}} A_x(b)e^{j\phi(b)}\left[\hat{g}[a(w-\phi_x'(b))] + \varepsilon(b,w)\right] \qquad 2\,4\,5$$

Where $g(\omega)$ is The Fourier transform of **g (t)** and $\varphi_x^1(t)$ is the time derivative of $\varphi(t)$. The term $\varepsilon(b, \omega)$ is an error term which is negligible if **A(t)** and $\varphi_x^1(t)$ have small variations over the support of $\varphi_{a,b}$, which is satisfied in most of the practical cases

Let ω_0 is the center frequency of the wavelet function $\varphi(t)$ and ω denote the center frequency of $\varphi_{a,b}$ Then frequency is related to the scale by

$$\omega = \frac{w_0}{a} \tag{2 4 6}$$

The normalized Scalogram is given by

$$\frac{\omega}{\omega_0} SC_x(b,\omega) = \frac{|T_x(b,a)|^2}{a} \quad for\ \omega = \frac{w_0}{a} \tag{2 4 7}$$

Using Eq (2 4 5), we get

$$\frac{\omega}{\omega_0} SC_x(b,\omega) = \frac{1}{4} A_s^2(b) \left| \hat{g}\left(\omega_0 \left[1 - \frac{\varphi_x'(b)}{\omega} \right] \right) + \varepsilon(b,\omega) \right| \tag{2 4 8}$$

Since $g(\omega)$ is maximum at $\omega=0$, neglecting $\varepsilon(b, \omega)$, the Scalogram is maximum at

$$\omega(b) = \frac{\omega_0}{a} = \varphi_x(b) \tag{2 4 9}$$

The corresponding points (b, ω(b)) are called wavelet ridges Thus ω(b) over the ridge gives the instantaneous frequency of the signal The instantaneous amplitude of the analytical signal is given by

$$A_s(b) = \frac{\sqrt{\frac{\omega}{\omega_0} SC_x(b,\omega)}}{|\hat{g}(0)|} \quad \text{along the ridge} \tag{2 4 10}$$

Thus ridge completely characterizes the analytical signal

2.5.Finding the ridge

Since |T (b, a)| is locally maximum on the ridge, the ridge can be tracked by finding the local maxima of the transform scale variable. Once the ridge is extracted the instantaneous frequency and amplitude can be extracted as follows

1 From the ridge, for each time b, corresponding to scale a, ω_0/a along the ridge represents the instantaneous frequency

2 The instantaneous amplitude can be extracted using Eq (2 4 10)

In the case of multi-component signals, if there is no overlap among the components, the following procedure is used to extract ridges

1 Compute the wavelet transform of analytical signal Find out the transform magnitude
2 Threshold the magnitude of the transform This suppresses the noise and cross terms when the regions are close
3 Find out different regions in the time-frequency plane, which represent the components in the multi-component signal.
4. Process each region masking the rest of the transform
5 Corresponding to each ridge, find out associated instantaneous frequency and amplitude

When the SNR is very low, the problem of choosing a proper threshold becomes critical, the threshold may miss a component that gives a peak below the threshold The above method is found to be robust for estimation of instantaneous frequency even in the presence of noise The presence of noise degrades the estimation of instantaneous amplitude

CHAPTER 3

Complex AM and FM Signal Models

3.1 Introduction

In real life, many signals, like speech, are non-stationary in nature Modelling of such non-stationary signals is rather a difficult task because of the inherent dependence of parameters of the model on time In many applications, one needs to represent a signal with a given model The parameters of the model are obtained with the signal in such a way that the synthetic version of the signal is as close to the original one as possible It is in this context, two novel methods are proposed to model non-stationary signals It has been found that voiced speech phonemes can be modelled using Complex AM signal model[3] whereas unvoiced speech phonemes can be modelled using Complex FM signal model[4] The estimation of the above model parameters is carried out using the Scalogram, the modulus of magnitude square of the continuous wavelet transform In this chapter each of the model is introduced and their suitability for modelling of speech phonemes is studied

3.2. Complex AM signal model

The complex sequence x [n] consisting of M single-tone amplitude modulated (AM) signals is represented by

$$x[n] = \sum_{i=1}^{M} A_i[1 + \mu_i e^{jv_i nT}]e^{jw_i nT} e^{j\phi} \qquad 3\ 2\ 1$$

Where

A_i is the carrier amplitude of constituent signal,

w_i is the carrier angular frequency,

μ_i is the modulation index

v_i is the modulating angular frequency and

ϕ_i is the independent and identically distributed(i i d) random phase which is assumed to be uniformly distributed over [0, 2π)

The signal model represented by (3 2 1) is highly non-stationary and estimating the model parameters is not an easy task The method of Scalogram developed in chapter 2 is used to estimate the parameters of the above model Once the carrier and modulating angular frequencies are determined from the time-frequency plot, the remaining problem of estimating amplitude and phases of each sub signal of the multi-component signal becomes a linear estimation problem as presented below

The discrete time signal x [n] fitted into the model of Eq (3 2 1) is reproduced below

$$x[n]=\sum_{i=1}^{M}A_i[1+\mu_i e^{jv_i nT}]e^{jw_i nT}e^{j\phi}$$

The above equation can be equivalently written as

$$x[n]=\sum_{i=1}^{M}A_{C_i}\xi_i^{\,n}+\sum_{i=1}^{M}A_{C_i}\mu_i\xi_i^{\,n}\zeta_i^{\,n},\quad n=0,1,2\ldots,N-1 \qquad 3\ 2\ 2$$

Where

A_{C_i} is the unknown complex amplitude of the carrier,

μ_i is the unknown modulation index, and

$\xi_i=e^{jw_iT},\zeta_i=e^{jv_iT}$ are the parameters which can be computed from the estimated parameters of carrier and modulating angular frequencies Rewriting Eq (3 2 2) in matrix form, we get

$$\begin{bmatrix} x[0]\\ x[1]\\ x[2]\\ .\\ .\\ .\\ x[N-1]\end{bmatrix}=\begin{bmatrix} 1 & . & 1 & 1 & . & 1\\ \xi_1 & .\ . & \xi_M & \xi_1\zeta_1 & . & \xi_M\zeta_M\\ \xi_1^{\,2} & & \xi_M^{\,2} & \xi_1^{\,2}\zeta_1^{\,2} & . & \xi_M^{\,2}\zeta_M^{\,2}\\ \\ \\ . \\ \xi_1^{\,N-1} & . & \xi_M^{\,N-1} & \xi_1^{\,N-1}\zeta_1^{\,N-1} & . & \xi_M^{\,N-1}\zeta_M^{\,N-1}\end{bmatrix}\begin{bmatrix} A_{c_1}\\ .\\ .\\ .\\ A_{c_M}\\ A_{c1}\mu_1\\ .\\ .\\ .\\ A_{c_M}\mu_M\end{bmatrix} \qquad 3\ 2.3$$

Eq (3 2 3) is to be solved in the least squares sense to find complex carrier amplitude and modulation index parameters The pseudoinverse of the matrix can be computed by a suitable decomposition technique like singular value decomposition (SVD) The estimation accuracy of carrier amplitude and modulation index parameters directly depends on the estimation accuracy of the carrier and modulating angular frequencies

3.2.1. Simulation Study

Consider the complex sequence {x [n], n=0, 1, 2 . N-1} consisting of two single tone amplitude modulated sub signals The parameters of the synthesized signal are as listed below

$A_1 = 1\ 0$, $w_1 = 0\ 35$, $\mu_1 = 0\ 30$, $v_1 = 0\ 005$, $\phi_1 = \pi/4$

$A_2 = 1\ 5$, $w_2 = 0\ 25$, $\mu_2 = 0\ 50$, $v_2 = 0\ 015$, $\phi_2 = \pi/6$

$A_3 = 1\ 0$, $w_3 = 0\ 15$, $\mu_3 = 0\ 00$, $v_3 = 0\ 010$, $\phi_3 = \pi/8$

The Scalogram of the above multi-component amplitude modulated signal is shown in Fig 3 1 One can clearly identify the carrier and modulating angular frequencies in the Scalogram plot The ridges of the Scalogram represent the carrier and carrier plus modulating angular frequencies, which were found to be in coincidence with the above values Also from the instantaneous amplitude plots, carrier amplitudes and modulation indices can be found Fig 3 3 represents the Scalogram of the above signal at 10 dB SNR and Fig 3 4 represents the Scalogram at 0 dB SNR It can be inferred that estimation of frequency parameters is robust to noise, whereas noise degrades the amplitude parameter estimation

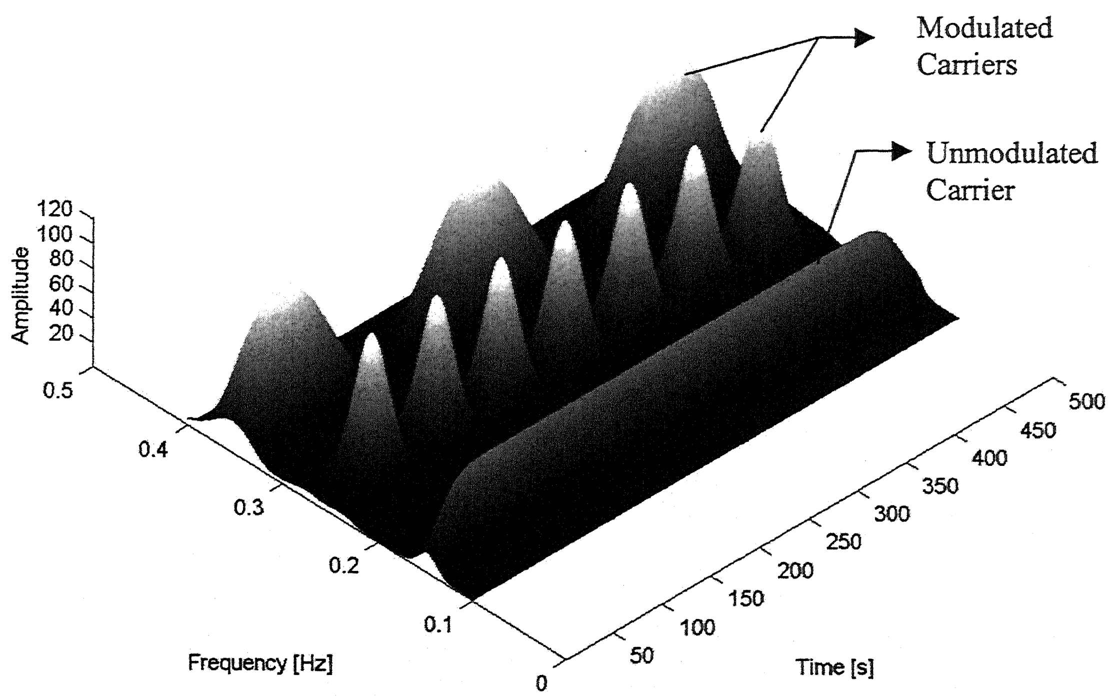

Fig. 3.1. Scalogram of 3 component AM Signal

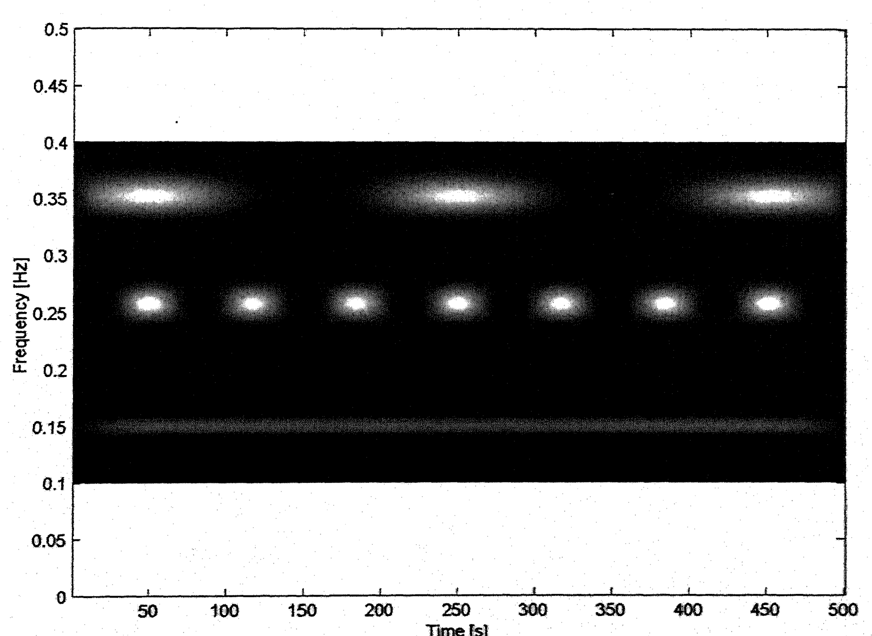

Fig. 3.2 Instataneous Frequency Plot

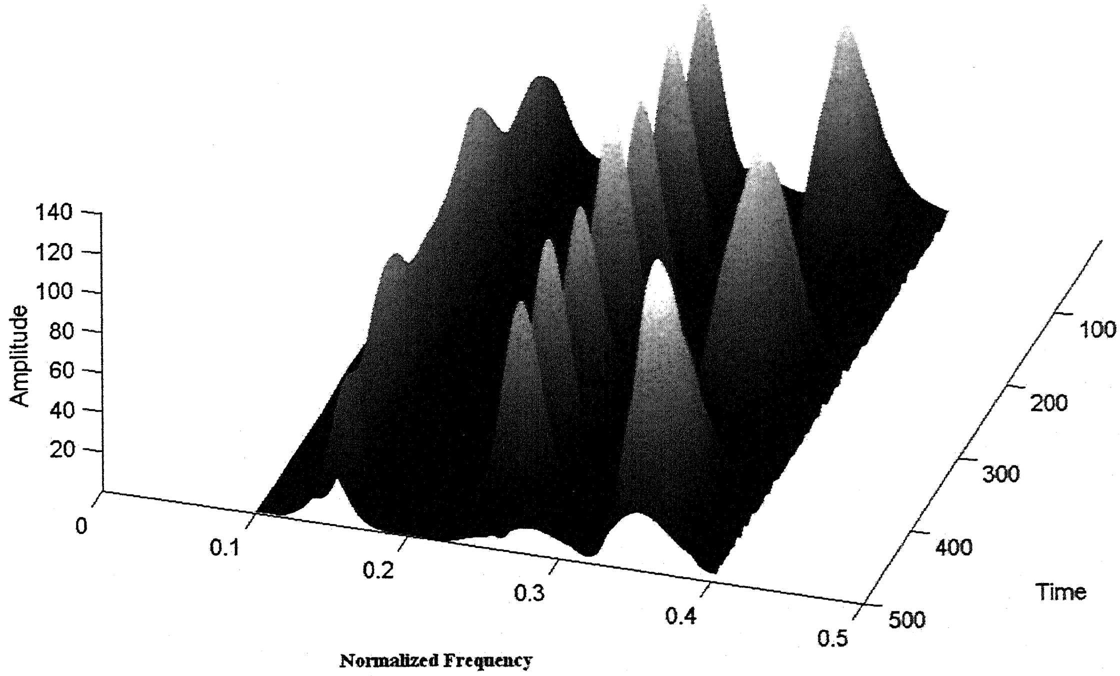

Fig 3.3 Scalogram of the 3 component AM signal at 10 dB SNR

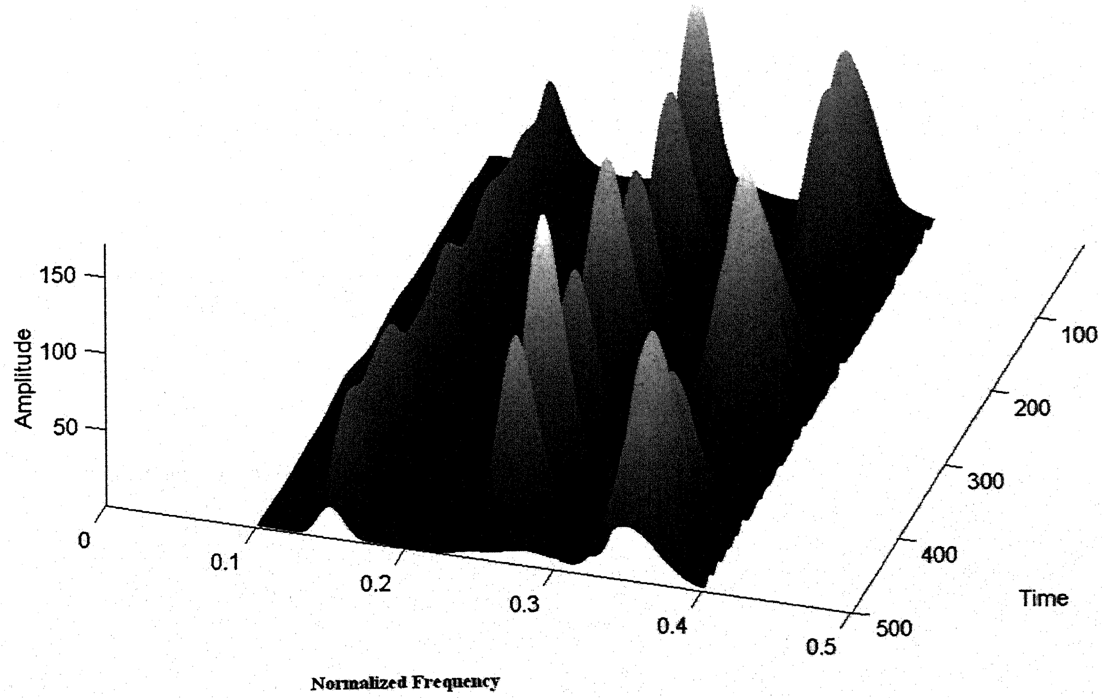

Fig. 3.4. Scalogram of the 3 component AM signal at 0 dB SNR

It has been shown that voiced speech phonemes can be modelled using complex AM signal model[3].

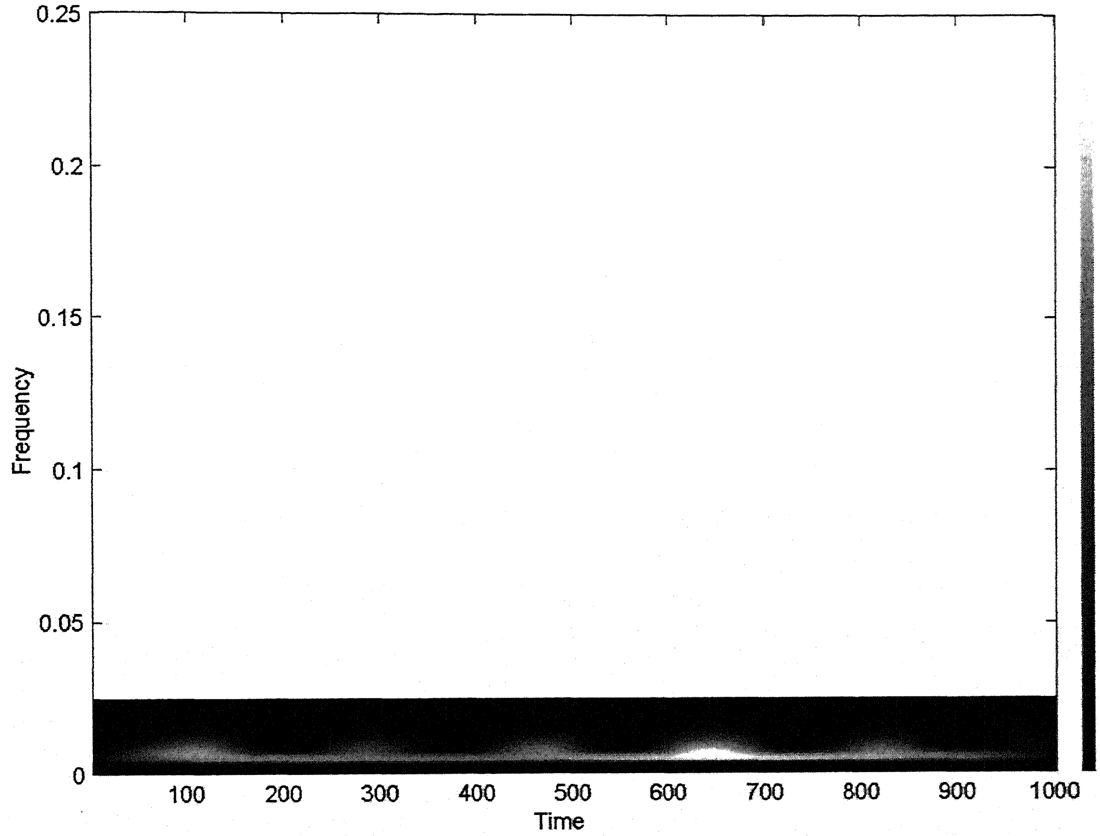

Fig. 3.5. Scalogram of Voiced Speech Phoneme /u/

Fig. 3.5 represents the Scalogram plot of voiced speech phoneme /u/. Looking at the plot, one can observe that it has only one amplitude modulated component. The original and reconstructed plots of speech phoneme are plotted in Fig 3.6.

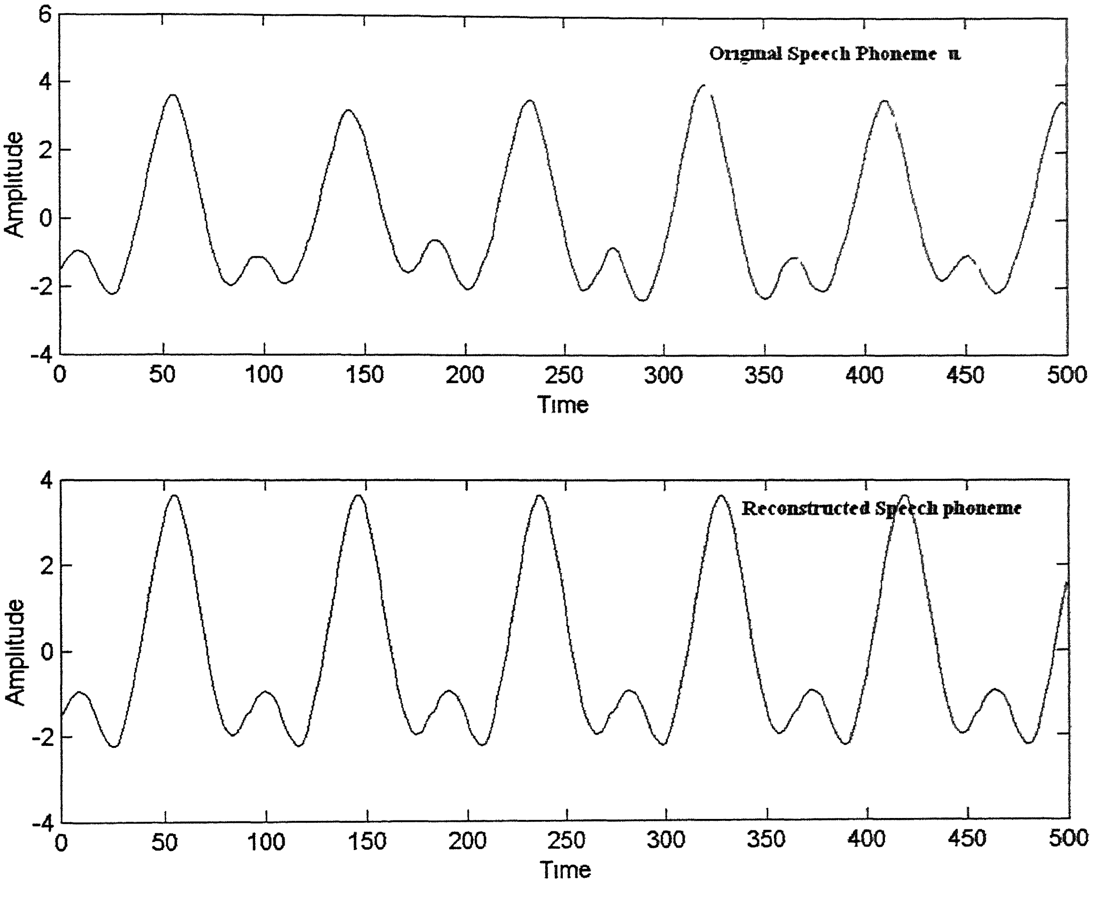

Fig. 3 6 Original and reconstructed plots of voiced speech phoneme /u/

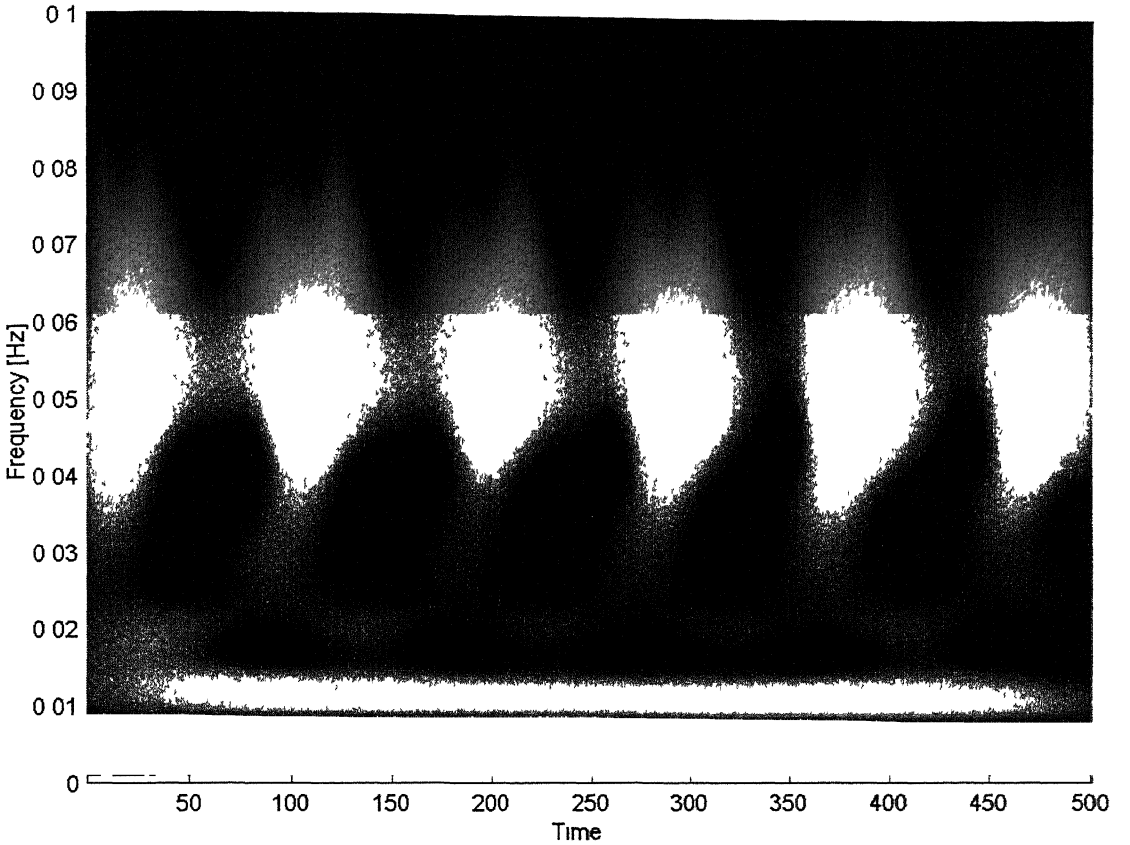

Fig 3 7. Scalogram of Voiced Speech Phoneme /a/

The above plot reveals that the voiced speech phoneme /a/ has two amplitude modulated carriers and one unmodulated carrier

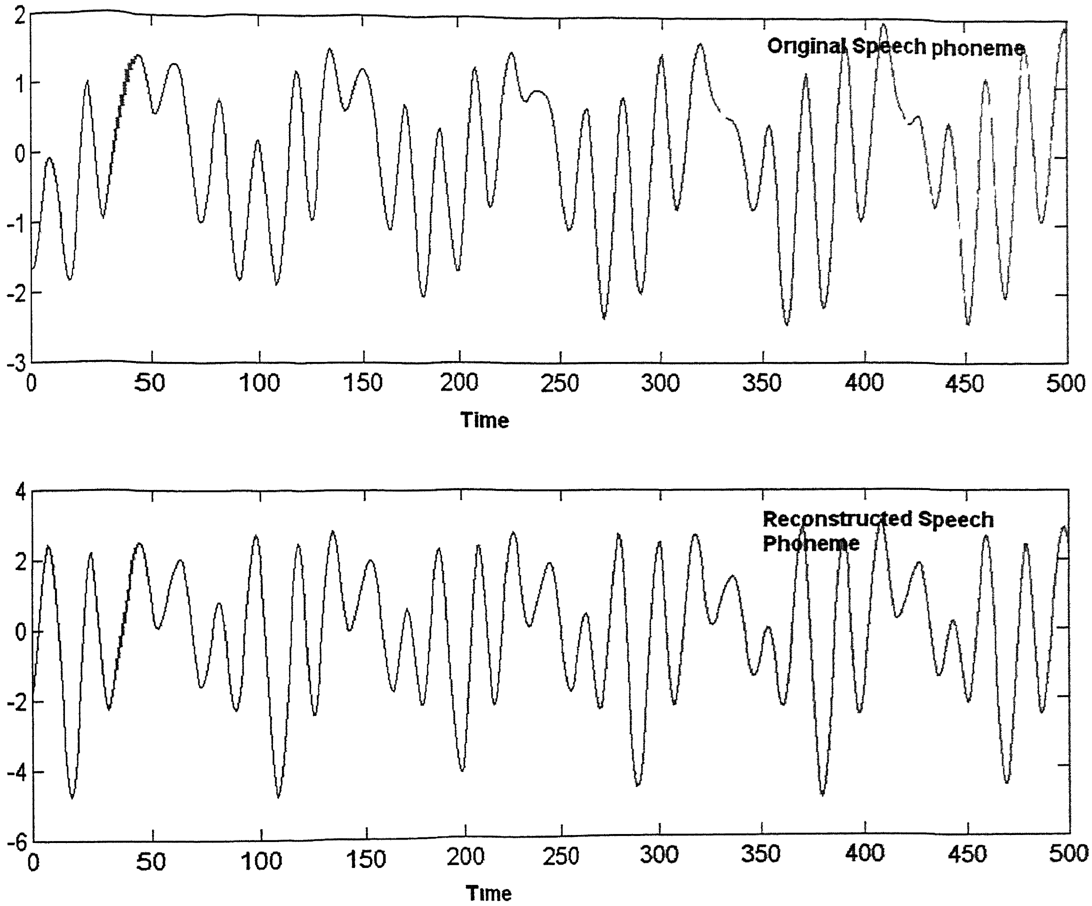

Fig 3.8. Original and Reconstructed Speech phoneme /α/

3.3. Complex FM signal model

The complex sequence x[n] consisting of M single tone frequency modulated (FM) subsequences can be represented as

$$x[n]=\sum_{i=1}^{M} A_i\, e^{j(w_i n+\phi_i+\beta_i \cos(v_i n))} \qquad 3\;3\;1$$

Where

A_i is the amplitude of complex exponential carrier signal

w_i is the carrier angular frequency,

β_i is the modulation index

v_i is the modulating angular frequency and

ϕ_i is the random phase independent and identically distributed(i i d) which is assumed to be uniformly distributed over $[0, 2\pi)$

The signal model given in Eq (3.3 1) is highly non-stationary and estimating the parameters of the model is not an easy task The method of Scalogram developed in chapter 2 will be used to estimate the parameters of the above model Once the carrier and modulating angular frequencies, modulation indices are determined from the Scalogram plot, the remaining problem of estimating the amplitudes and phases of each sub signal of the multi-component signal becomes a linear estimation problem

Consider the complex sequence {x [n], n=0, 1, 2, N-1} given by Eq. (3 3 1) which is reproduced below

$$x[n]=\sum_{i=1}^{M} A_i \, e^{j(w_i n+\phi_i+\beta_i \cos(v_i n))}$$

This can be equivalently written as

$$x[n]=\sum_{i=1}^{M} A_{C_i} W_{n_i} \qquad 3\,3\,2$$

Where the complex amplitudes and the complex weights given below can be computed with the estimated parameters of carrier, modulating angular frequencies and modulation indices

$$\left.\begin{aligned} A_{C_i} &= A_i e^{j\phi_i} \\ W_{n_i} &= e^{j(w_i n+\beta_i \cos(v_i n))} \end{aligned}\right\} \qquad 3\,3\,3$$

Rewriting (3 3.2) in matrix form, we get

$$\begin{bmatrix} x[0] \\ x[1] \\ \cdot \\ \\ \cdot \\ x[N-1] \end{bmatrix} = \begin{bmatrix} W_{0,1} & W_{0,2} & \cdots & W_{0,M} \\ W_{11,} & W_{1,2} & \cdot\ \cdot & W_{1,M} \\ \cdot & & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ W_{N-1,1} & W_{N-1,2} & \cdot\ \cdot & W_{N-1,M} \end{bmatrix} \begin{bmatrix} A_{c_1} \\ A_{c2} \\ \cdot \\ \cdot \\ \cdot \\ A_{cM} \end{bmatrix} \qquad 3\,3\,4$$

Eq (3 3 4) should be solved in least squares sense by employing a suitable pseudoinverse of the complex weight matrix which will be over determined with N > M Note that the problem of solving Eq (3 3 5) depends directly on the estimation accuracy of the angular and modulation index parameters

3.3.1. Simulation Study

Consider the complex sequence {x [n], n = 0, 1, 2 N} consisting of two single tone frequency modulated sub signals The parameters of the synthesized signal parameters are as listed below

$$A_1 = 1\ 0,\ w_1 = 0\ 35,\ \beta_1 = 0\ 10,\ v_1 = 0\ 02,\ \phi_1 = \pi/4$$

$$A_2 = 1\ 5,\ w_2 = 0\ 15,\ \beta 2 = 0.05,\ v_2 = 0\ 01,\ \phi 2 = \pi/6$$

The Scalogram of the above multi-component frequency modulated signal signal is shown in Fig 3 9 Clearly one can observe how the frequency is varying in the Scalogram plot Fig 3 10 represents instantaneous frequency of individual components recovered from the ridges of the Scalogram The values of angular, modulating frequencies as well as modulation indices can be estimated from the ridges of the Scalogram, which were found to be in coincidence with the above values Now with the help of Eq (3 3 5), the complex amplitudes can be calculated Fig 3 11 represents the Scalogram of the above signal at 10 dB SNR Fig3.12 represents the Scalogram at 0 dB SNR It can be inferred that the effect of noise is minimal on the estimation of frequency parameters where as amplitude estimation degrades heavily with the presence of noise

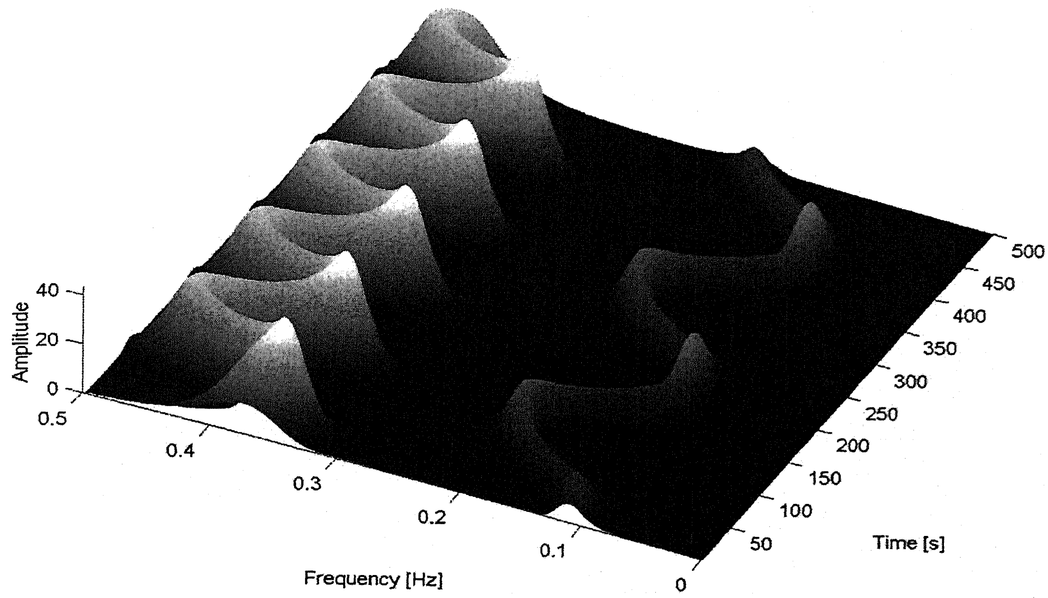

Fig. 3.9. Scalogram of two component FM signal

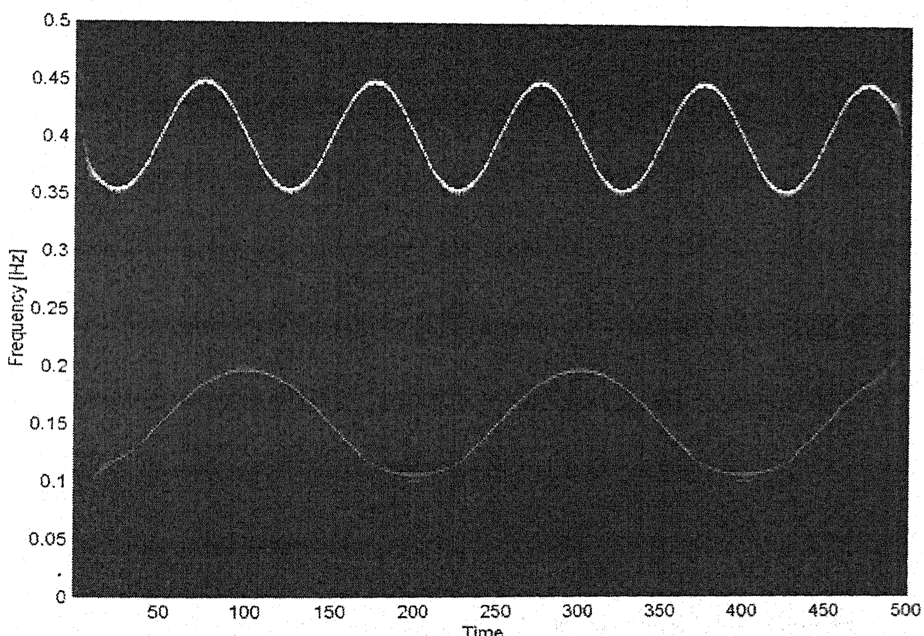

Fig. 3.10. Ridges of the Scalogram

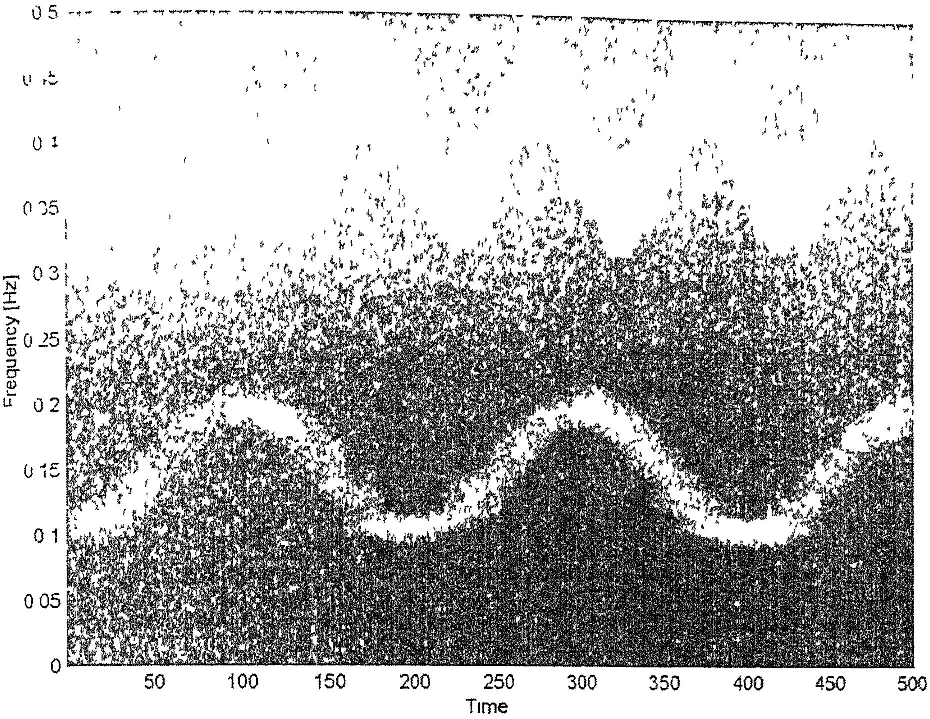

Fig 3 11 Scalogram of two component FM signal at 10 dB SNR

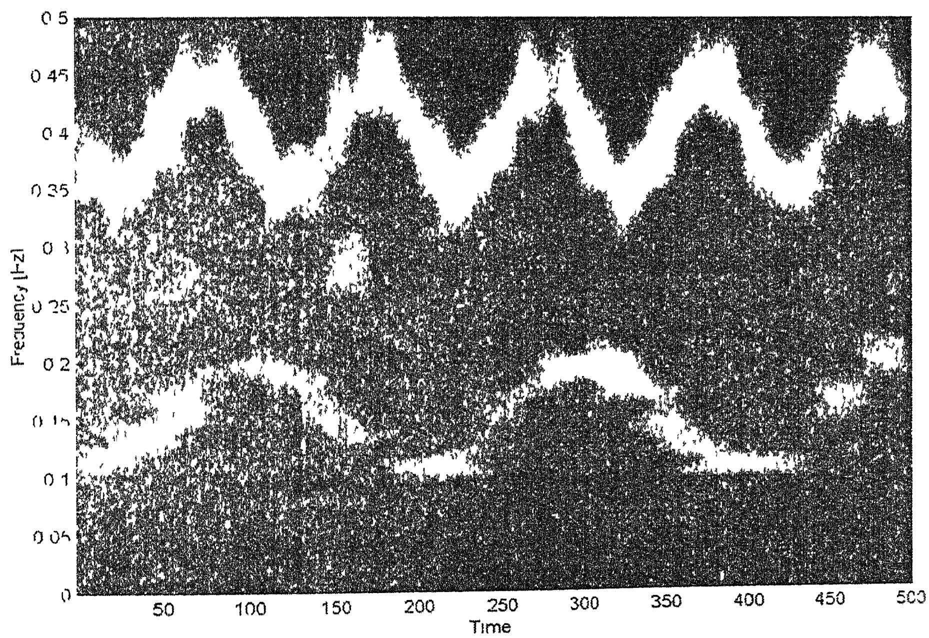

Fig 3 12 Scalogram of two component FM signal at 0 dB SNR

It has been shown[4] that unvoiced phonemes can be modelled using complex FM model.

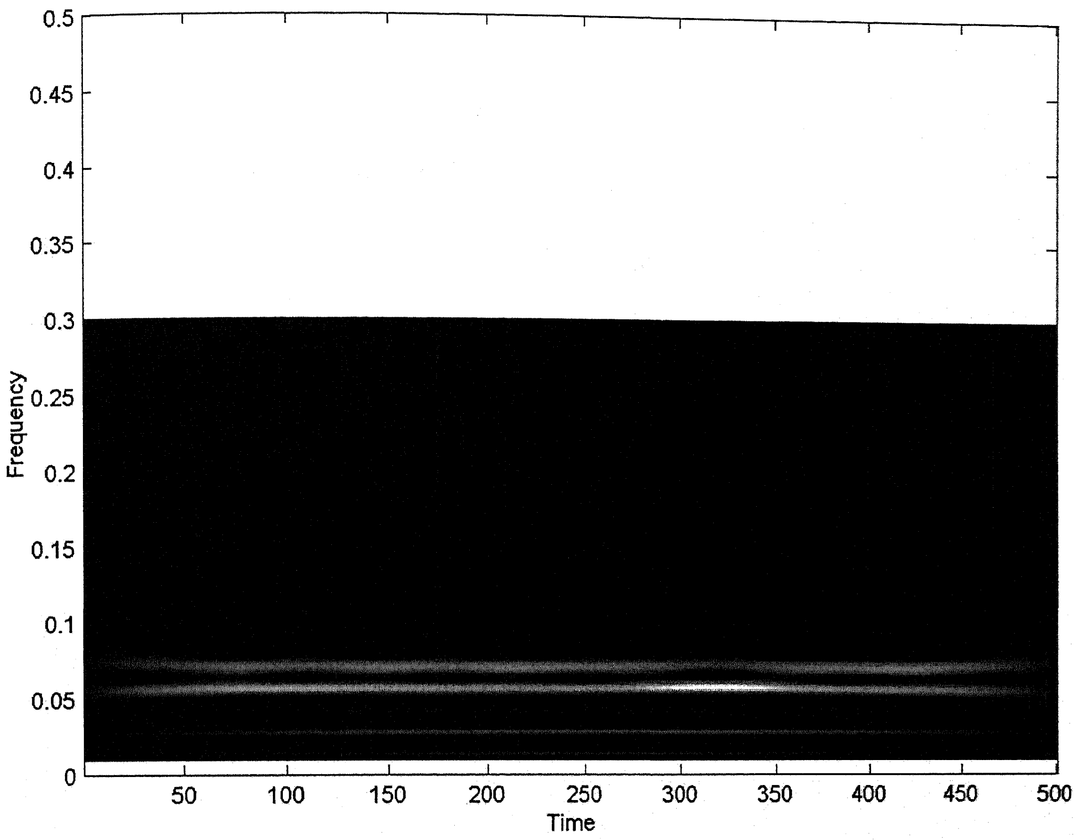

Fig 3.13. Scalogram of unvoiced speech phoneme /q/

From the scalogram of Fig. 3.13, one can clearly observe that the unvoiced speech phoneme /q/ has two frequency modulated components and three unmodulated carriers. These parameters can be estimated from the instantaneous frequency plots, which were obtained from the ridges of the scalogram.

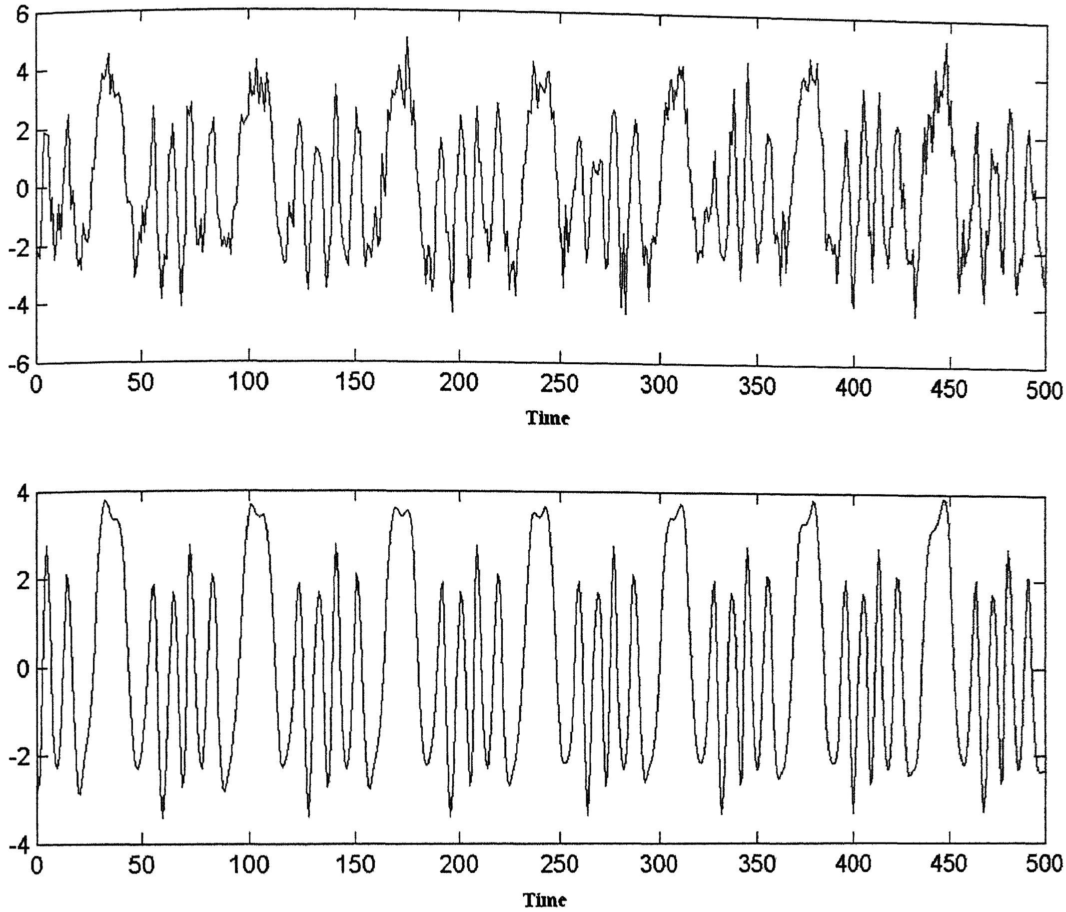

Fig 3 14 (a) Original speech phoneme /q/ (b) Reconstructed speech phoneme /q/

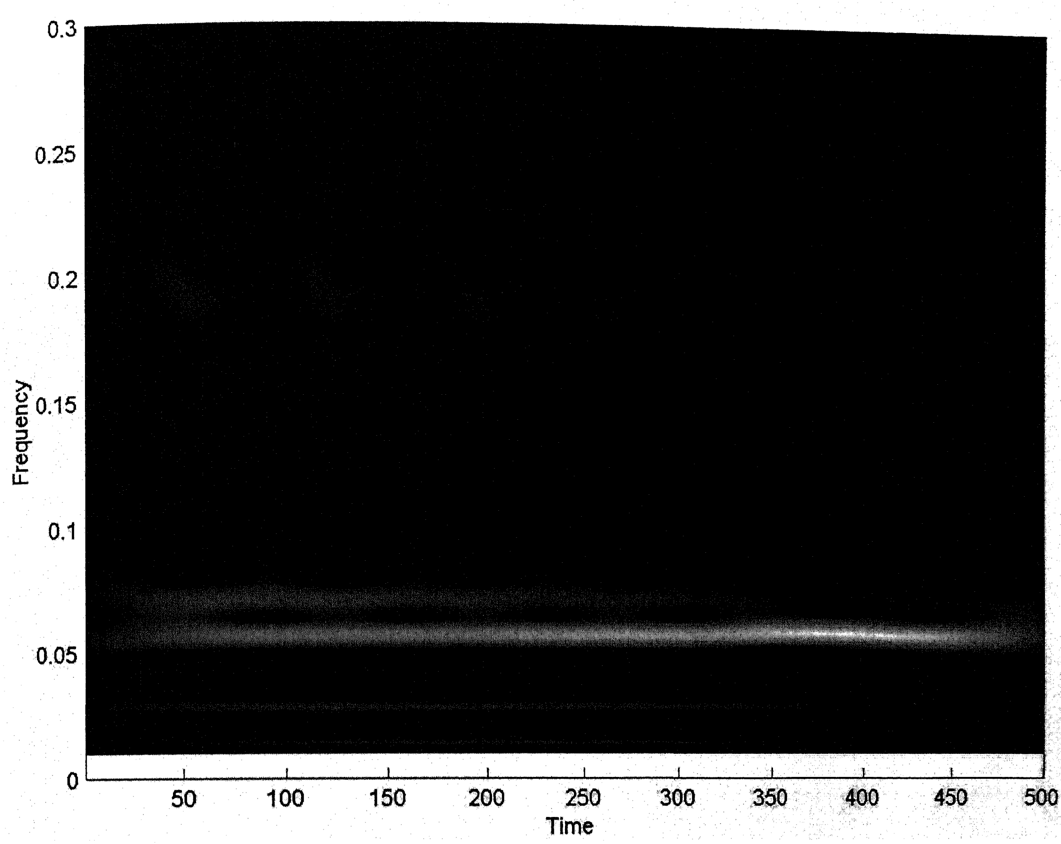

Fig 3.15 Scalogram of speech phoneme /v/

From the scalogram of Fig. 3.15, one can clearly observe that the unvoiced speech phoneme /q/ has three frequency modulated components and three unmodulated carriers. These parameters can be estimated from the instantaneous frequency plots, which were obtained from the ridges of the scalogram.

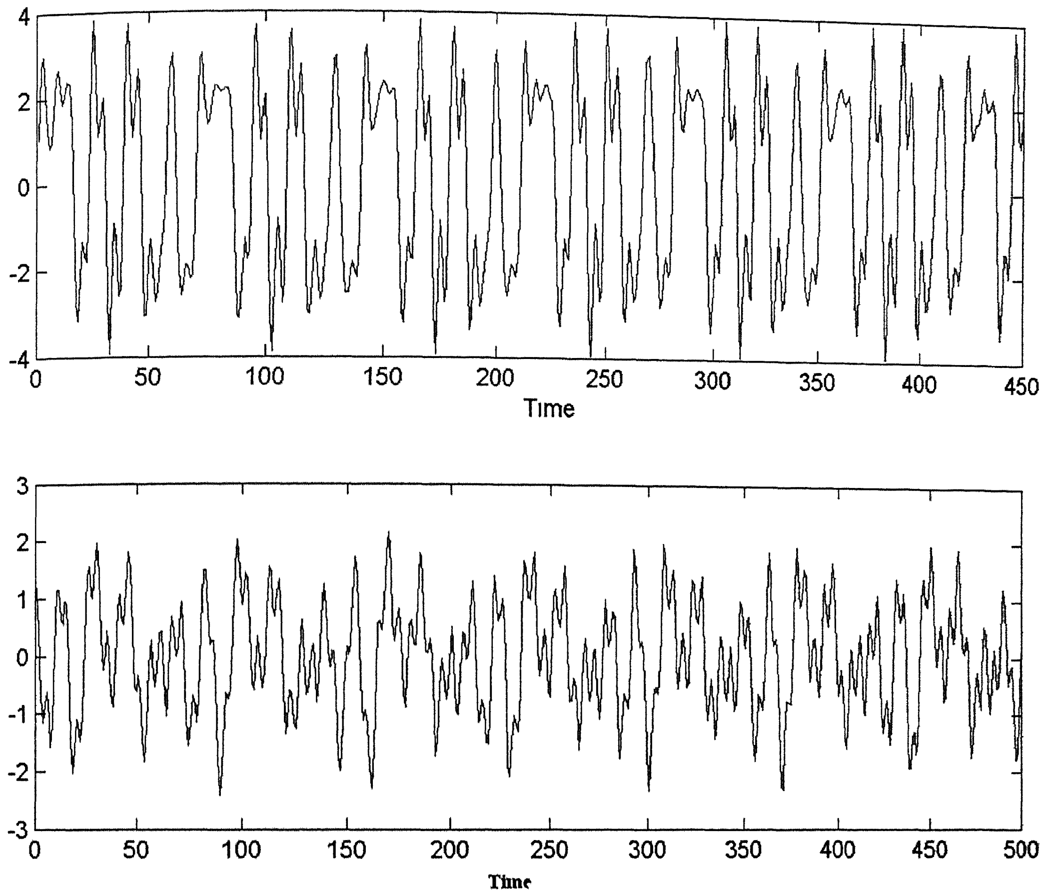

Fig 3 16(a) Original speech phoneme /v/ (b) Reconstructed speech phoneme /v/

Chapter 4

CONCLUSIONS

In this dissertation time frequency distributions were reviewed and scalogram, a time frequency distribution of affine class has been extensively studied Morlet wavelet has been used to estimate the time varying frequency and amplitude quantities of the multi-component signals It has been found that the ridges of the scalogram completely characterize the components present in a multi-component signal, since the time frequency representation tend to localize in disjoint regions in the time frequency plane

Complex AM signal model has been found to be quite suitable for modelling voiced speech phonemes whereas Complex FM signal model has been found to be suitable for modelling unvoiced speech phonemes The simulations revealed that Scalogram technique is a powerful tool to analyze multi-component signals This method fails when the components present in the multi-component model overlap with each other Also it is has been found that the frequency estimation is robust to noise while the amplitude estimates are sensitive to noise

SUGGESTIONS FOR FURTHER WORK

In the present work it has been established that voiced speech phonemes can be modelled using complex AM signal model, whereas unvoiced speech phonemes can be modelled using complex FM signal model Hence a combination of these two models can be used to model a continuous speech data Moreover the Scalogram technique fails when there is significant overlap among different components in the time-frequency plane In that case, time frequency distributions that give more resolution in quadratic time-frequency plane are to be employed for separating the components

Bibliography

[1] L Cohen, "Tıme-frequency dıstrıbutıons—A revıew," *Proc IEEE*, Vol. 77, pp 941–981, July 1989

[2]. B Boashash, "Estımatıng and ınterpretıng the ınstantaneous frequency of a sıgnal," *Proc IEEE*, part I, vol 80, pp 520–538, part II, Vol 80, pp 540–568, 1992

[3] P Sırcar, M S Syalı "Complex AM sıgnal model for non statıonary sıgnals" *Sıgnal Processıng* 57(1996) pp 35-45

[4] P Sırcar, Sanjay Sharma "Complex FM sıgnal model for non statıonary sıgnals" *Sıgnal Processıng* 57(1997) pp 283-304

[5] H I Choı and W J Wıllıams, "Improved time-frequency representatıon of multıcomponent sıgnals usıng the exponentıal kernels," *IEEE Trans Acoust, Speech, Sıgnal Processıng*, vol 37, pp 862–871, June 1989

[6] N Delprat *et al*, "Asymptotıc wavelet and Gabor analysıs Extractıon of ınstantaneous frequencıes," *IEEE Trans Inform Theory*, Vol 38, pp 644–664, 1992

[7] Leman Helene, Marque Catherıne "Rıdge Extractıon from the scalogram of the uterıne electromyogram" *IEEE Trans Sıgnal Processıng pp 245-248*

[8] R Carmona, W L Hwang, and Torr´esanı "Multırıdge Detectıon and Tıme–Frequency Reconstructıon" *IEEE Trans On Sıgnal Processıng, Vol 47, No 2,1999*

[9] R Carmona, W L Hwang, and B Torr´esanı, "Characterızatıon of sıgnals by the rıdges of theır wavelet transforms," *IEEE Trans Sıgnal Processıng*, vol 45, pp 2586–2590, 1997

[10] Thuong Le-Tion "Some issues of wavelet functions for instantaneous frequency estimation" *IEEE TENCON 1997Speech and Imaging technologies for Computing and Telecommunications,* p p 31-34

[11] S Mallat and S Zhong, "Characterization of signals from multiscale edges," *IEEE Trans Pattern Anal Machine Intell* vol 14, July 1992

[12] Patrick Fladrin, Jochim Stockler "Time Frequency/Time Scale Analysis" Academic Press, 1998